www.icgst.com

Improved Face Recognition based on Hidden Markov Model

Sameh Magdy, Tarek Mahmoud, Mohamed Ibrahim
School of Communication Engineering, Canadian international College
Sameh_mohamed@cic-cairo.com,
https://www.cic-cairo.com/cic/

Abstract

In this paper, a new face recognition technique based on Hidden Markov Model (HMM), Pre-processing, and feature extraction (K-means and the Sobel operator) is proposed. Two main contributions are presented; the first contribution in the pre-processing were image's edges are normalized to enhance the HMM models to be non-sensitive to different edges. The second contribution is a new technique to extract the image's features by splitting the image into non-uniform height depending on the distribution of the foreground pixels. The foreground pixels are extracting by using the vertical sliding windows. The proposed technique is faster with a higher accuracy with respect to other techniques which are investigated for comparison. Moreover, it shows the capability of recognizing the normal face (center part) as well as face boundary.

Keywords: *HMM, Sobel operator, Face Recognition , Accuracy*.

Nomenclature:

DCT	Discrete Cosine Transform
HMM	Hidden Markov Model
LDA	Linear discriminate Analysis
ORL	Olivetti Research Laboratory
PCA	Principal Component Analysis
PDBNN	Probabilistic Decision- Based Neural Network
SVD	Singular Values Decomposition
W_o	represents the width of overlapping of the vertical frames

1. Introduction

Face recognition is identifying a human face among other different faces. It is widely used in all recent applications particularly related to the security field. For example, patterns recognition techniques are very essential in the treatment of the facial images to identify the criminal act, ID and credit card verification. Face recognition has become necessary in commercial and law enforcement researches. Recently, dedicated work targets face recognition process, which is aiming to handle the trade-off between time and accuracy. However, face recognition faces too many challenges such as the difference between human faces in color, shape, and dimensions. Therefore, there is a need for image analysis to extract its landmarks and forms, and then convert it into data to be processed. There are two main categories used for face detection; the first category is the feature based where analysis is performed on the main geometry of the face like eyes, nose, ears, and mouth. The second category is a holistic base where the face is analyzed as a two-dimensional pattern[1]. A wide range of models can be used for addressing this issue, most commonly models are Geometrical Feature Matching [2], Eigenfaces [3], Neural Networks [4], Linear discriminate analysis (LDA) [5], Principal component analysis (PCA), Discrete cosine transform (DCT) [6] and the fuzzy logic technique [7], which is used for face detection and Hidden Markov Models. Geometrical feature matching model is based on extracting geometrical features to form a complete picture of the face. This model showed a low accuracy with 75% on ORL database recognition rate. Another model is Eigenfaces method which uses the Principal Component Analysis (PCA) with 90.5% recognition rate on ORL database. Although this model is simple and fast however it is not robust enough over the changes in face orientation and lighting conditions. A neural network is a widely used model in many pattern recognition applications. The paper [8] used Probabilistic Decision- Based Neural Network (PDBNN) for face recognition with 96 % recognition rate. The recognition problems of this model appear when increasing faces classes. Here a used stochastic process is the hidden Markov model. Hidden Markov Model (HMM) is used for developing features extraction methods and enhancing not only the operation of face detection but also the data processing time as will be shown in this paper. In this paper, we introduce a new method for pre-processing the input image and a new method of extracting features by splitting the input image horizontally with non-uniform heights. This paper is organized as follows. In section 2, we explain our proposed technique then, we present the pre-processing step for the input image in section A. In section B we explain the process of the feature extraction. In section C, we introduce the use of HMM in the recognition system. In section 3 we give a full description of the obtained simulation results followed by the conclusions and future work is highlighted in section 4.

2. The Proposed Technique

Our proposed technique consists of multiple steps:

- The pre-processing step which will be explained in the following section.
- The feature extraction to prepare data to be processed.
- Appling HMM on the recognition system.

We calculate the accuracy (recognition rate) to measure the ability of recognize the image by the system.Figure1 shows the block diagram for our proposed technique.

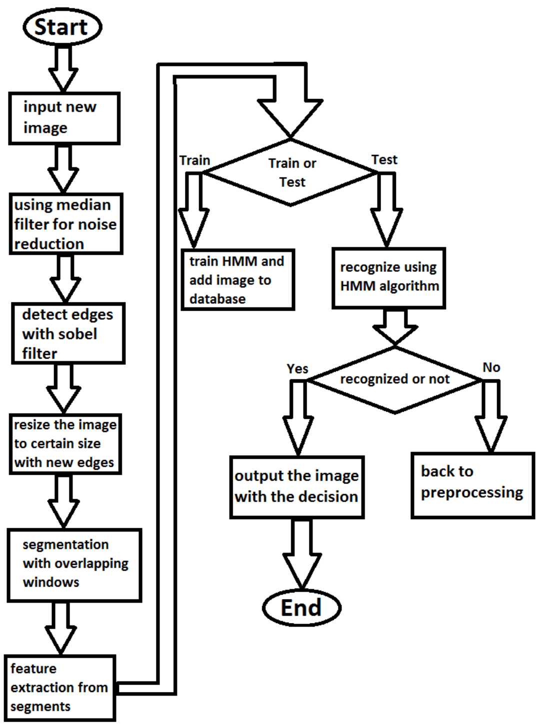


Figure 1 Face Recognition Proposed Technique

Pseudo code

```
READ new input image
        not_known = 1
WHILE (not_known)
    DO median filter to reduce noise
    DO other enhancement
    DO sobel filter to detect horizontal and vertical edges

IF (image_size! = threshold size)
            DO resize the image to new size
  ENDIF

DO segmentation of the resized image with overlapping
sliding window (horizontal and vertical)
DO extract features from the segments of the image
DO KMEANS and Clustering
            IF (HMM_training)
                    IF (new_class)
                        DO add new class model
                    ELSE
        DO modify HMM class model using the image
                    ENDIF
          ELSE
          Classify the image using the HMM models
                    IF recognized
          WRITE output the image with decision
                              Break from while loop
                    ENDIF
          ENDIF
ENDWHILE
```

A. Pre-processing

The first step is to change the image from RGB format to its analogous grayscale format. A grayscale image is a monochrome image, which contains information about the brightness only without containing any information about the color. A grayscale data represents a wide range of intensities. The typical image contains 8 bit/pixel, this allows the image to be represented with different level of brightness (0-255) [9]. The pre-processing is an essential step to enhance the ease of features extraction and to reduce the noise in the image by using the median filter which is used to remove impulsive noise and decrease edges blurring. Pre-processing stage is important step in enhancing the performance of the whole system due to the following reason:

- normalizing the image's edges to enhance the

HMM models and makes it not sensitive to different edges. The recognition rate of our system without the pre-processing stage is 97%, after pre-processing the recognition rate 100% (enhanced 3 %).

B. Feature Extraction

After the pre-processing stage, the input image in binary format in the ORL database is used to extract a feature vector, which labeled the background pixels by logic "0" and the foreground pixels by logic "1". The input image is segmented into n horizontal segments have the same number of foreground pixels in each one of them (segment). The total number of pixels in the image is counted and then divided by n which results in the required number of pixels per segment. Then the input image is divided into vertical overlapping frames. Each one of these frames is divided into fixed width small cells, while the cell's height depends on the pixel's image and the distribution of foreground pixels in the images. The sliding window is shifted along the image from right to left and a feature vector is calculated for each frame, as shown in Figure2.

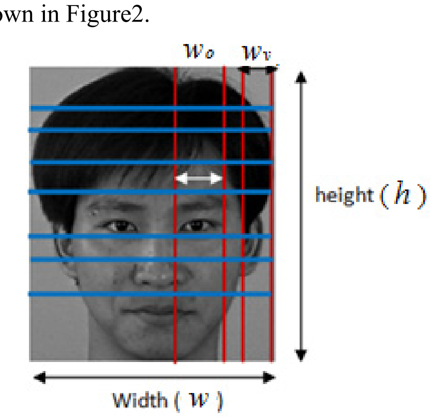


Figure.2 Feature extraction using a sliding window with overlapping.

The image's height represents (h), the image's width represents (w), (w_o) represents the width of overlapping of the vertical frames, and (w_v) represents the width of the vertical frame. The best scenario in our technique when ' w_v '= 7 pixels with ' w_o '= 4 pixels (trial and error), this sliding accomplished the 100% accuracy (recognition rate) as it extracted the maximum necessary features for recognition. 6 cells (trial and error) are used per vertical frame with non uniform heights to extract this feature for each cell, the number of "1" pixels are calculated to produce 6 features per frame (6 cells per frame x1 feature per cell) as shown in figure3.

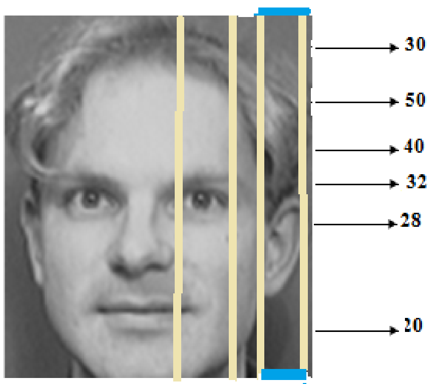


Figure 3 Six features per frame by dividing the image into six non uniform heights.

- ## Gradient Features

The gradient is a vector which has certain magnitude and direction.

$$\nabla f = \begin{pmatrix} \frac{\delta f}{\delta x} \\ \frac{\delta f}{\delta y} \end{pmatrix}$$

<u>Advantage of the sobel operator:</u>

I. Faster to compute gradient operator because Gaussian smoothing effect.

II. Less sensitive to noise present in image and smoothing affects the accuracy of edge detection.

The magnitude of gradiant provide information about the strength of the edges[10].

$$\nabla f = \sqrt{(\frac{\delta f}{\delta x})^2 + (\frac{\delta f}{\delta y})^2} = \sqrt{(M_x)^2 + (M_y)^2} \qquad \textbf{(1)}$$

The Direction is per-pendicular to the direction of the edge $\nabla f = \tan^{-1}(M_y / M_x)$ **(2)**

This is done by applying the Sobel operator on the image to extract the vertical and horizontal gradient components [6]. To compute the feature vector, only the direction is used, which can range from 0 to 360 degrees. This range is divided into 8 non-overlapping regions each of them is 45 degree. In each sampling region at each pixel, a histogram of gradient directions is taken. In each region step, histogram value corresponds to the count of each gradient direction. To extract the gradient features we used two cells with non-uniform heights per frame (trial and error). For each cell, eight gradient features are produced to form 16 features per frame (2 non-uniform cells per frame x 8 features per cell) as shown in figure 4

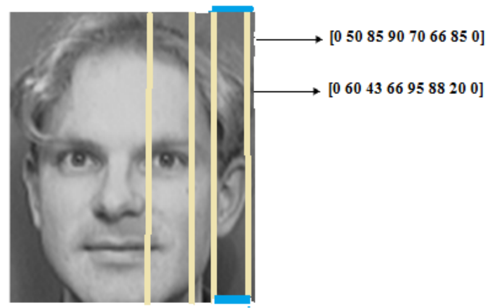


Figure 4 Sixteen feature per frame

<u>*Pseudo-codes for Sobel edge detection method*</u>

Input: A Sample Image

Output: Detected Edges

Step1: Accept the input image

Step2: Apply mask M_x, M_y to the input image

Step3: Apply Sobel edge detection algorithm and the gradient

Step 4: Masks manipulation of M_x, M_y separately on the input image

Step 5: Results combined to find the absolute magnitude of the gradient

Step 6: the absolute magnitude is the output edges

- ## Clustering and K-mean

Clustering technique is a well known tech. in the field of pattern recognition, image and data analysis. The concept of clustering can be defined as a process of creating a group of sample, each group contains number of sample which are similar, so this groups are called clusters. Many techniques of clustering are proposed using the hierarchical for clusters of a unit element [11]. Considering the two main factors (simplicity and efficiency), clustering technique are used for the natural image segmentation. In partitional clustering, our objective is to create one set of cluster, where data is divided among similar groups. In our work, K-means clustering approach is used to perform image segmentation using Matlab software. We use clustering to classify the set pixels that have same characteristics to be in one cluster. K-means clustering is easy to be applied in large datasets [12]. First we made to define one K-centroid for each cluster, which is considered the main idea. Then, when the centroid are organized in a cunning way as different locations well cause different results, so it is better to place centroid far away with respect to each other. After that each point related to the given data set is associated to the nearest centroid [13]. For each step a new K-centroids are required to be recalculated. Then a new binding has to be done between the nearest new centroid and the same data set points. A loop has been generated. In this paper, our results show that the k-centroid change their location in a different

way until the k-centroid have a defined fixed location as shown in table1.

__Pseudo-codes for K-means and clustering__

X: a set of N data vectors Data set

C_I: initialized k cluster centroids Number of cluster

C: the cluster centroids of k-clustering random initial centroids

$P = \{p(i) \mid i = 1, \ldots, N\}$ is the cluster label of X

KMEANS(X, C_I) → (C, P)

REPEAT

$C_{previous} \leftarrow C_I$;

FOR all $i \in [1, N]$ DO

Generate new optimal paritions

$p(i) \leftarrow \arg\min_{l \le j \le k} d(x_i, c_j)$;

FOR all $j \in [1, k]$ DO

Generate optimal centroids

$c_j \leftarrow$ Average of x_i, whose $p(i) = j$;

UNTIL $C = C_{previous}$

C. Hidden Markov Model

HMM is a complex stochastic process, which needs to be done two times to assure the generation of an efficient model for the sequential data.HMM is a defined set of states, and transition probabilities control the transitions among those states. HMM process deals with all feature observations in all states and then produce probability distribution to indicate the confidence of the object represented by the model. The classification can be obtained from an unknown pattern Based on the extracted feature sequence by finding a model that generates a maximum probability [4].The HMM used in this paper is a discrete HMM, the used HMM classifier is based on the same classifier implemented in HTK Recognition Toolkit [14]. It uses the Viterbi algorithm, which searches for the most likely sequence of a hidden state given the input feature vector. Three states HMM are depicted in figure5, which shows that we allowed a transition to the current and the next states only. Two of these are emitting states and have their output probability distributions. The transition matrix for this model consists of 3 rows and 3 columns, which represent the transitions between these states such that incase of transition, we will set the matrix element to one and matrix element will be set to zero in case of no transition. The element of HMM are:

- N is the number of state in the model. If S is the set of state, then $S = \{S_1, S_2, S_3, \ldots S_N\}$. The state of the model at time t is given by $q_t \in S, 1 \le t \le T$, where T is the length of the observation sequence (number of frames).
- M, the number of different observation symbols. If V is the set of all possible observation symbols $V = \{v_1, v_2, \ldots, v_M\}$.
- A the state of transition probability matrix i.e. $A = \{a_{ij}\}$ where

$$a_{ij} = P[q_t = S_j \setminus q_{t-1} = S_i] \quad (3)$$

$1 \le i, j \le N$ with the constraint , $0 \le a_{i,j} \le 1$ and,

$$\sum_{j=1}^{N} a_{ij} = 1, \qquad 1 \le i \le N$$

- B, the observation of symbol probability matrix i.e. $B = \{b_j(k)\}$, where

$$b_j(k) = P[O_t = v_k \setminus q_t = S_j] \quad (4)$$

$1 \le j \le N, 1 \le K \le M$ and O_t is the observation symbol at time t.

- π, the initial state distribution, i.e. $\prod = \{\pi_i\}$ where

$$\pi_i = P[q_1 = S_i], 1 \le i \le N \quad (5)$$

Using a notation, a HMM is defined as the triplet

$$\lambda = (A, B, \pi) \quad (6)$$

the above characterization corresponds to a discrete HMM. In a continuous density HMM, the state are characterization by continuous observation density functions. The most general representation of the model probability density function (PDF) is a finite mixture of the form:

$$b_i(o) = \sum_{k=1}^{M} C_{ik} N(O, \mu_{ik}, U_{ik}) \quad (7)$$

where, C_{ik} is the mixture coefficient for the Kth mixture in state i. without losses of generality $N(O, \mu_{ik}, U_{ik})$ is assumed to be a Gaussian PDF with mean vector μ_{ik} and covariance U_{ik}

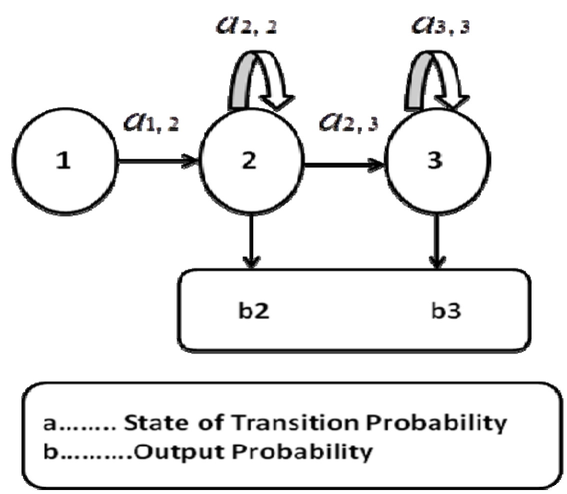


Figure 5 Three states (transition) left to right

3. Simulation Results

To evaluate the performance of the proposed recognition technique, recognition trials have been conducted on available subsets of the benchmark ORL face database. In our technique validation, we used 80 images for training purpose (8 persons × 10 images), and 40 images for testing purpose (8 persons × 5 images). The face recognition technique presented in this paper was developed, trained, and tested using Matlab 7.10.The computer was a Windows 7 machine with a 2.1 GHz Intel core 2 processor and 2.87 GB of RAM. The following parameters have been determined through validation: Number of states, and the number of chosen centroids. The following table 1 shows the effect of changing centroids and stats versus time. The following fig.6 shows samples of ORL database. Figure 7 and 8 shows the effect of the given data in table1. Table 2 and 3 shows the comparison ORL database with different technique in Recognition rate and training time. Figure .9,10 and shows the effect of the given data in table 2and 3. Figure11 shows the Comparison Results of Different Face recognition algorithms (Testing stage).

Figure 6 Sample of ORL face database.

Table1 Shows the effect of varying the number of centroids and stats on the results.

Model	**Recognition Rate (%)**	**Training Time/ image**
Proposed Technique	**100%**	**0.00329 sec**
PDBNN	96%	20 min
n-tuple	86%	0.9
1DHMM + Wavelet	100%	1.13 sec
Pseudo-2D HMM	95%	1 sec
DCT-HMM	99.5 %	23.5 sec
1D HMM + SVD	99 %	0.63 sec

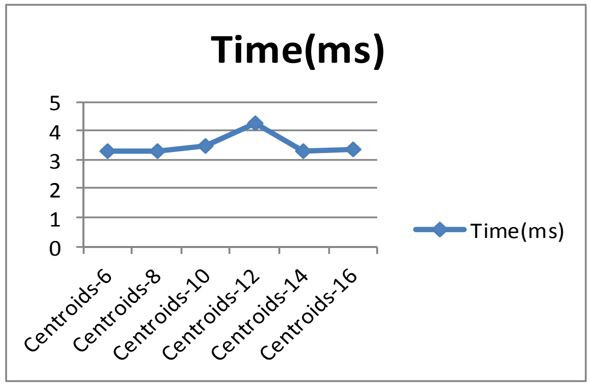


Figure 7 Time (ms) versus the number of centroids

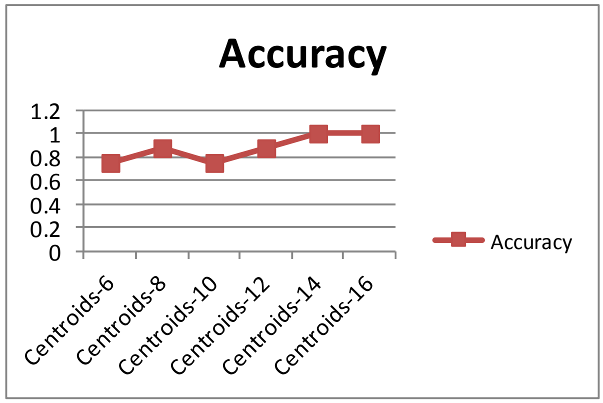


Figure.8 Accuracy versus the number of centroids.

Table 2 Comparison ORL database with different technique in Recognition rate and training time.

#Centroids	Average Time/Image(sec)	Recognition rate (Accuracy %)
6	0.00326	75
8	0.00329	87.5
10	0.00346	75
12	0.00427	87.5
14	0.00329	100
16	0.00336	100

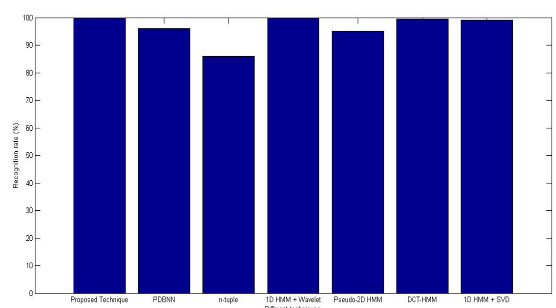


Figure 9 Comparison Results of Different Face recognition algorithms and proposed technique.

The recognition rate of the proposed technique in this paper is higher than that of the traditional algorithms based on HMM. Also, we compare our proposed technique with different face recognition methods such as DCT-HMM, FisherFace, Laplacianfaces, ML-HMM, MCE-HMM,Pseudo-2D HMM,PDBNN and MC-HMM [15], our proposed method gives the highest recognition accuracy for all the sample sets. Comparison result is summarized in the following Table2 and figure. 9 and figure10.

TABLE 3 COMPARISON RESULTS OF DIFFERENT FACE RECOGNITION ALGORITHM

Models	Recognition rate (%)
Proposed Technique	**100**
E-HMM	99.16
Eigenfaces	85.9
Fisherfaces	92.3
Laplacianfaces	93.2
ML-HMM	88.2
MCE-HMM	93.3
MC-HMM	97.5

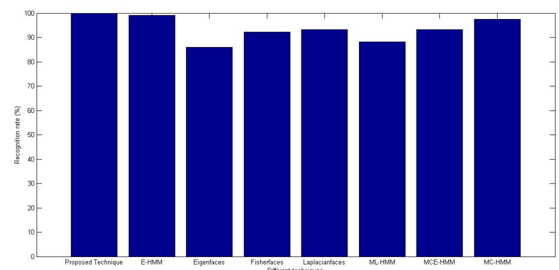


Figure 10 Comparison Results of Different Face recognition algorithms and proposed technique.

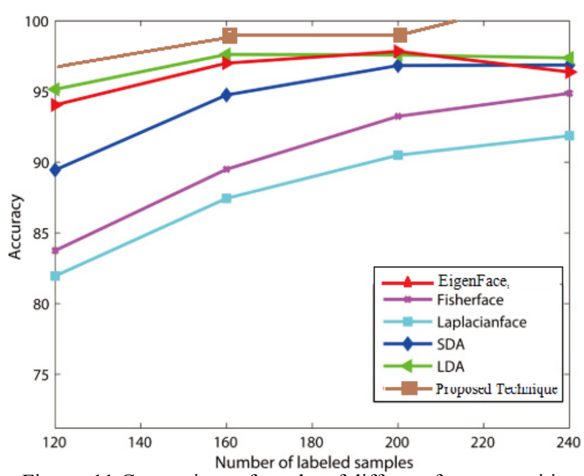


Figure 11 Comparison of results of different face recognition algorithms (Testing stage).

It is worth mentioning the technique has been implemented without performing the pre-processing step mentioned in section A. The results for four different sets of databases showed that the recognition rate is about 97 %. This implies clearly the effect of the pre-processing step on the recognition rate.

4. Conclusion

In this paper, we introduced a new effective technique for image face recognition. We proposed our technique based on intelligent features using efficient operator (Sobel). The First main contribution pre-processing stage has improved the performance of the overall system by about 3 % by normalizing the image's edges to enhance the HMM models and makes it not sensitive to different edges. Although k-means, sobel and pre-processing are important steps, they simplify the features and enhance the recognition rate as shown in results. The second main contribution is a new technique of splitting the image into non uniform heights which has improved the performance of the system, when compared with using fixed heights. Our proposed technique showed the best performance compared with the other tested techniques in the sense of speed and recognition rate. .

Acknowledgments

We would like to express the deepest gratitude and sincere appreciation to Al-Azhar University and Canadian university for offering their help, support, and encouragement in accomplishing this paper.

5. References

[1] Xavier P Burgos-Artizzu, Pietro Perona, and Piotr Dollár. “Robust face landmark estimation under occlusion. In Proc”. IEEE Int. Conf. Computer. Vision., pages 1513–1520, 2013.

[2] Gurpreet Kaur, Manbir Sandhu and Purnima.” Facial Recognition: Issues, Techniques and Applications”International Journal of Advanced Research in Computer Science and Software Engineering Volume 6, Issue 2, February 2016, pp. 508-512

[3] [Navneet Jindal , Vikas Kumar, “ Enhanced Face Recognition Algorithm using PCA with Artificial Neural Networks”, Electronics and Communication Engineering Department Samalkha Group of Institutes, IJARCSSE, pp 864-872, 2013.

[4] Manisha M. Kasar , Debnath Bhattacharyya and Taihoon Kim.” Face Recognition Using Neural Network: A Review” International Journal of Security and Its Applications Vol. 10, No. 3 (2016), pp.81-100

[5] Suman kumar Bhattacharyya and Kumar rahul.”Face Recognition by Linear Discriminant Analysis ” International Journal of Communication Network Security, ISSN: 2231- 1882, Volume-2, Issue-2, 2013

[6] P. Phaneemdra, V. Ramachandran\and E. Srinivasa Reddy.“Human Face Detection and Recognition using PCA” and DCT in HMM. International Journal of scientific Engineering and technology research ISSN 2319- 8885 Vol.04, Issue.35, Pages: 7080-7085, August-2015.

[7] Shweta Mehta, Shailender Gupta, Bharat Bhushan.” Face recognition using Neuro- Fuzzy Interface System ”International Journal of Signal Processing, Image Processing and Pattern Recognition.Vol.7,No.1.July 2014 PP.339-344.

[8] Nisha, Sandeep Dahiya. “Face Detection and Expression Recognition using Neural Network Approaches”. Global Journal of Computer Science and Technology: F Graphics & Vision Volume 15 Issue 3 Versions 1.0 Year 2015.

[9] Hany Ahmed, Sherif Abdel Azeem. “Effective Technique for the Recognition of Write Independent Off-line. Handwritten Arabic words”. In proceedings of 13th International Conference on Frontiers in Handwriting, Bari, Italy, September 2012.

[10] Mrs. Swati Y.Raut, Mrs.Dipti.A.Doshi. “A face recognition system by Hidden Markov Model and Discriminating set approach”. International Journal of Scientific Engineering Research Volume 6, Issue.1, January-2015 30 ISSN 2229-5518.

[11] Chien, J-T. & Liao, C-P. “Maximum Confidence Hidden Markov Modeling for Face Recognition. Pattern Analysis and Machine Intelligence”. IEEE Transactions Vol. 30, No 4, pp. 606-616, April 2008.

[12] N. Senthilkumaran, R. Rajesh. “Edge Detection Techniques for Image Segmentation and A Survey of Soft Computing Approaches”. International Journal of Recent Trend in Engineering, Vol. 1, No. 2, PP.250-254, May 2009.

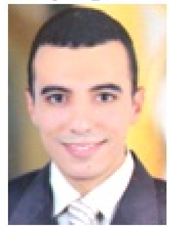
[13] WenshuoGao, et.al. “An improved Sobel edge detection”. Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference, China, Volume: 5, pp. 67 – 71, 9-11 July 2010.

[14] Tapas Kanungo, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman. “An Efficient k-Means Clustering Algorithm: Analysis and Implementation” IEEE transactions on pattern analysis and machine intelligence, vol. 24, no. 7, july 2002

[15] T. Kanungo, D.M. Mount, N.S. Netanyahu, C. Piatko, R. Silverman, and A.Y. Wu. “The Analysis of a Simple k- Means Clustering Algorithm”, Proc. 16th Ann. ACM Symp. Computational Geometry, pp 100-109, June 2000.

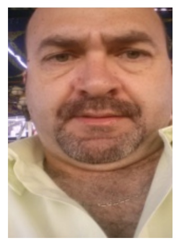
[16] Mrs. Swati Y.Raut ,and Mrs.Dipti.A.Doshi. “A Face Recognition System by Hidden Markov Model and Discriminating Set Approach”. International Journal of Scientific &Engineering Research, Volume 6, Issue 1, January-2015.

Biographies



Sameh Abaza B.sc degree in Electronics and communication from 10th of Ramadan University 2005 .Finished master degree from Al-Azhar University 2009, Cairo, Egypt, in satellite image resolution enhancement technique and works as a lecturer in Electrical and communication department at Canadian international university. Email: sameh_mohamed@cic-cairo.com , Tele: +20-01006551609

Prof. Mohamed Ibrahim was born 1949. Finished Ph.D. jan1988 West Germany Ruhr University. He worked with Prof. Alfred, a famous Prof in Digital signal processing field .Worked Now as a Prof in Electrical and communication Engineer at Al-Azhar University, Egypt. He worked in Al-Azhar University from 1974. Published 174 research in different fields (Signal and image processing, Digital Filter and Algorithms for digital signal processing, Digital Communication). Email: mohiyosof@yahoo.com.



Tarek Mahmoud Ph.D. of electrical engineering, ECE department, University of Victoria, 2003, BC, Canada., M.Sc. of electrical engineering, Al-Azhar University,
1997, Cairo, Egypt. B.Sc. of electrical engineering. Ain- shams University, 1990, Cairo, Egypt. Distinction with First class honor. Published 4 researches in IEEE International Symposium on signal processing and Information Technology, Email: Tarekmahmoud@hotmail.com